

S. Kaushal

7/1/01
9/27/01 OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,493

DATE: 07/18/2001

TIME: 10:37:01

Input Set : A:\96700-613.2.ST25.txt

Output Set: N:\CRF3\07182001\I516493.raw

ENTERED

3 <110> APPLICANT: Charron, Maureen
 4 Katz, Ellen
 6 <120> TITLE OF INVENTION: NOVEL GLUCOSE TRANSPORTER/SENSOR PROTEIN AND USES THEREOF
 8 <130> FILE REFERENCE: 96700/613
 10 <140> CURRENT APPLICATION NUMBER: US 09/516,493
 12 <141> CURRENT FILING DATE: 2000-03-01
 14 <150> PRIOR APPLICATION NUMBER: US 09/356,602
 16 <151> PRIOR FILING DATE: 1999-07-19
 18 <160> NUMBER OF SEQ ID NOS: 12
 20 <170> SOFTWARE: PatentIn version 3.0
 23 <210> SEQ ID NO: 1
 25 <211> LENGTH: 98
 27 <212> TYPE: PRT
 29 <213> ORGANISM: mouse
 31 <400> SEQUENCE: 1
 33 Ile Val Ala Ile Phe Gly Phe Val Ala Phe Phe Glu Ile Gly Pro Gly
 34 1 5 10 15
 36 Pro Ile Pro Trp Phe Val Ala Glu Leu Phe Ser Gln Gly Pro Arg Pro
 37 20 25 30
 39 Ala Ala Met Ala Val Arg Gly Phe Ser Asn Trp Thr Cys Asn Phe Ile
 40 35 40 45
 42 Val Gly Met Gly Phe Gln Tyr Val Ala Asp Arg Met Gly Pro Tyr Val
 43 50 55 60
 45 Phe Leu Leu Phe Ala Val Leu Leu Leu Gly Phe Phe Ile Phe Thr Phe
 46 65 70 75 80
 48 Leu Lys Val Pro Glu Thr Arg Gly Arg Thr Phe Asp Gln Ile Ser Ala
 49 85 90 95
 51 Ala Phe
 54 <210> SEQ ID NO: 2
 56 <211> LENGTH: 100
 58 <212> TYPE: PRT
 60 <213> ORGANISM: mouse
 62 <400> SEQUENCE: 2
 64 Val Gly Ser Met Cys Leu Phe Ile Ala Gly Phe Arg Val Gly Trp Gly
 65 1 5 10 15
 67 Pro Ile Pro Trp Leu Leu Met Ser Glu Ile Phe Pro Leu His Ile Lys
 68 20 25 30
 70 Gly Val Ala Thr Gly Val Cys Val Leu Thr Asn Trp Phe Met Ala Phe
 71 35 40 45
 73 Leu Val Thr Lys Glu Phe Asn Ser Ile Met Glu Ile Leu Arg Pro Tyr
 74 50 55 60
 76 Gly Ala Phe Trp Leu Thr Ala Ala Phe Cys Ile Leu Ser Val Leu Phe
 77 65 70 75 80
 79 Thr Leu Thr Phe Val Pro Glu Thr Lys Gly Arg Thr Leu Glu Gln Ile
 80 85 90 95
 82 Thr Ala His Phe
 84 100

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,493

DATE: 07/18/2001

TIME: 10:37:01

Input Set : A:\96700-613.2.ST25.txt

Output Set: N:\CRF3\07182001\I516493.raw

```

87 <210> SEQ ID NO: 3
89 <211> LENGTH: 50
91 <212> TYPE: PRT
93 <213> ORGANISM: Saccharomyces
95 <400> SEQUENCE: 3
97 Ile Ala Phe Ile Cys Leu Phe Ile Ala Ala Phe Ser Ala Thr Trp Gly
98 1          5          10          15
100 Gly Val Val Trp Val Val Ser Ala Glu Leu Tyr Pro Leu Gly Val Arg
101          20          25          30
103 Ser Lys Cys Thr Ala Ile Cys Ala Ala Ala Asn Trp Leu Val Asn Phe
104          35          40          45
106 Thr Cys
107      50
110 <210> SEQ ID NO: 4
112 <211> LENGTH: 50
114 <212> TYPE: PRT
116 <213> ORGANISM: Saccharomyces
118 <400> SEQUENCE: 4
120 Ile Ala Phe Ile Cys Leu Phe Ile Ala Ala Phe Ser Ala Thr Trp Gly
121 1          5          10          15
123 Gly Val Val Trp Val Ile Ser Ala Glu Leu Tyr Pro Leu Gly Val Arg
124          20          25          30
126 Ser Lys Cys Thr Ala Ile Cys Ala Ala Ala Asn Trp Leu Val Asn Phe
127          35          40          45
129 Ile Cys
130      50
133 <210> SEQ ID NO: 5
135 <211> LENGTH: 50
137 <212> TYPE: PRT
139 <213> ORGANISM: Saccharomyces
141 <400> SEQUENCE: 5
143 Val Gly Ser Met Cys Leu Phe Ile Ala Gly Phe Arg Val Gly Trp Gly
144 1          5          10          15
146 Pro Ile Pro Trp Leu Leu Met Ser Glu Ile Phe Pro Leu His Ile Lys
147          20          25          30
149 Gly Val Ala Thr Gly Val Cys Val Leu Thr Asn Trp Phe Met Ala Phe
150          35          40          45
152 Leu Val
153      50
156 <210> SEQ ID NO: 6
158 <211> LENGTH: 1813
160 <212> TYPE: DNA
162 <213> ORGANISM: homo sapiens
164 <220> FEATURE:
166 <221> NAME/KEY: Unsure
168 <222> LOCATION: (1697)..(1697)
170 <223> OTHER INFORMATION: 'n' may be any one of a, t, c, or g
172 <400> SEQUENCE: 6
174 aacttgccgc cgccgcgtct tcctcgccgc cttcgccgct gccctgggcc cactcagctt      60

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,493

DATE: 07/18/2001

TIME: 10:37:01

Input Set : A:\96700-613.2.ST25.txt

Output Set: N:\CRF3\07182001\I516493.raw

```

176 cggcttcgcg ctcggtacac gctccccggc catccctagc ctgcagcgcg ccgcgcccc 120
178 ggccccgcgc ctggacgacg ccgcgcctc ctggttcggg gctgtcgtga ccctgggtgc 180
180 cgcggcgggg ggagtgtgg gcggtggct ggtggaccgc gccgggcgca agctgagcct 240
182 cttgctgtgc tccgtgccct tegtggccgg ctttgccgtc atcaccgcgg ccaggagct 300
184 gtggatgctg ctggggggcc gcttcctcac cggcctggcc tgcgggtgtg cctccctagt 360
186 ggccccggtc tacatctccg aaatcgccca cccagcagtc cgggggttgc tcggctcctg 420
188 tgtgcagcta atggtcgtcg tcggcatcct cctggcctac ctggcaggct ggggtcgtga 480
190 gtggcgctgg ctggtgtgtg tgggtgtgct gccccctcc ctcatgtgc ttctcatgtg 540
192 cttcatgccc gagacccgcg gcttcctgct gactcagcac aggcgccagg aggccatgc 600
194 cctgcggttc ctgtggggct ccgagcaggg ctgggaagac ccccccatcg gggctgagca 660
196 gagctttcac ctggccctgc tgcggcagcc cggcatctac aagcccttca tcacgggtgt 720
198 ctccctgatg gccttcagc agctgtcggg ggtcaacgcc gtcatgttct atgcagagac 780
200 catctttgaa gaggccaagt tcaaggacag cagcctggcc tcggtcgtcg tgggtgtcat 840
202 ccagggtgctg ttcacagctg tggcggtctt catcatggac agagcagggc ggaggctgct 900
204 cctggtcttg tcagggtgtg tcatggtgtt cagcacgagt gccttcggcg cctacttcaa 960
206 gctgaccagc ggtggccctg gcaactcctc gcacgtggcc atctcggcg ctgtctctgc 1020
208 acagcctggt gatgccagcg tggggtggc ctggtggcc gtgggcagca tgtgcctctt 1080
210 catgcggcg tttgcgggtg gctgggggc catccctgg ctctcatgt cagagatctt 1140
212 ccctctgcat gtcaagggcg tggcgacagg catctgcgtc ctcaccaact ggctcatggc 1200
214 ctttctcgtg accaaggagt tcagcagcct catggaggte ctcaggccct atggagcctt 1260
216 ctggcttgcc tccgctttct gcatcttcag tgtcctttc actttgttct gtgtccctga 1320
218 aactaaagga aagactctgg aacaaatcac agccatttt gagggcgcat gacagccact 1380
220 cactagggga tggagcaagc ctgtgactcc aagctgggcc caagccaga gccctgcct 1440
222 gccccagggg agccagaatc cagcccttg gagccttgg ctgcagggtc cctccttct 1500
224 gtcatgtctc ctccagccca tgaccgggg ctaggaggct cactgcctcc tgttccagct 1560
226 cctgctgctg ctctgaggac tcaggaacac cttcagactt tgcagacctg cggtcagccc 1620
228 tccatgcgca agactaaagc agcggaagag gaggtgggcc tctaggatct ttgtcttctg 1680
w 230 gctggagggt cttttgnagg ttgggtgctg ggcattcggg cgctcctctc acgggctgc 1740
232 cttatcgga aggaatttg tttgccaat aaagacgtga cacagaaaat caaaaaaaaa 1800
234 aaaaaaaaaat tcc 1813

```

237 <210> SEQ ID NO: 7

239 <211> LENGTH: 453

241 <212> TYPE: PRT

243 <213> ORGANISM: homo sapiens

245 <400> SEQUENCE: 7

```

247 Arg Arg Val Phe Leu Ala Ala Phe Ala Ala Ala Leu Gly Pro Leu Ser
248 1 5 10 15
250 Phe Gly Phe Ala Leu Gly Tyr Ser Ser Pro Ala Ile Pro Ser Leu Gln
251 20 25 30
253 Arg Ala Ala Pro Pro Ala Pro Arg Leu Asp Asp Ala Ala Ala Ser Trp
254 35 40 45
256 Phe Gly Ala Val Val Thr Leu Gly Ala Ala Ala Gly Gly Val Leu Gly
257 50 55 60
259 Gly Trp Leu Val Asp Arg Ala Gly Arg Lys Leu Ser Leu Leu Leu Cys
260 65 70 75 80
262 Ser Val Pro Phe Val Ala Gly Phe Ala Val Ile Thr Ala Ala Gln Asp
263 85 90 95
265 Val Trp Met Leu Leu Gly Gly Arg Leu Leu Thr Gly Leu Ala Cys Gly
266 100 105 110

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,493

DATE: 07/18/2001

TIME: 10:37:01

Input Set : A:\96700-613.2.ST25.txt

Output Set: N:\CRF3\07182001\I516493.raw

```

268 Val Ala Ser Leu Val Ala Pro Val Tyr Ile Ser Glu Ile Ala Tyr Pro
269      115      120      125
271 Ala Val Arg Gly Leu Leu Gly Ser Cys Val Gln Leu Met Val Val Val
272      130      135      140
274 Gly Ile Leu Leu Ala Tyr Leu Ala Gly Trp Val Leu Glu Trp Arg Trp
275 145      150      155      160
277 Leu Ala Val Leu Gly Cys Val Pro Pro Ser Leu Met Leu Leu Leu Met
278      165      170      175
280 Cys Phe Met Pro Glu Thr Pro Arg Phe Leu Leu Thr Gln His Arg Arg
281      180      185      190
283 Gln Glu Ala Ile Ala Leu Arg Phe Leu Trp Gly Ser Glu Gln Gly Trp
284      195      200      205
286 Glu Asp Pro Pro Ile Gly Ala Glu Gln Ser Phe His Leu Ala Leu Leu
287      210      215      220
289 Arg Gln Pro Gly Ile Tyr Lys Pro Phe Ile Ile Gly Val Ser Leu Met
290 225      230      235      240
292 Ala Phe Gln Gln Leu Ser Gly Val Asn Ala Val Met Phe Tyr Ala Glu
293      245      250      255
295 Thr Ile Phe Glu Glu Ala Lys Phe Lys Asp Ser Ser Leu Ala Ser Val
296      260      265      270
298 Val Val Gly Val Ile Gln Val Leu Phe Thr Ala Val Ala Ala Leu Ile
299      275      280      285
301 Met Asp Arg Ala Gly Arg Arg Leu Leu Leu Val Leu Ser Gly Val Val
302      290      295      300
304 Met Val Phe Ser Thr Ser Ala Phe Gly Ala Tyr Phe Lys Leu Thr Gln
305 305      310      315      320
307 Gly Gly Pro Gly Asn Ser Ser His Val Ala Ile Ser Ala Pro Val Ser
308      325      330      335
310 Ala Gln Pro Val Asp Ala Ser Val Gly Leu Ala Trp Leu Ala Val Gly
311      340      345      350
313 Ser Met Cys Leu Phe Ile Ala Gly Phe Ala Val Gly Trp Gly Pro Ile
314      355      360      365
316 Pro Trp Leu Leu Met Ser Glu Ile Phe Pro Leu His Val Lys Gly Val
317      370      375      380
319 Ala Thr Gly Ile Cys Val Leu Thr Asn Trp Leu Met Ala Phe Leu Val
320 385      390      395      400
322 Thr Lys Glu Phe Ser Ser Leu Met Glu Val Leu Arg Pro Tyr Gly Ala
323      405      410      415
326 Phe Trp Leu Ala Ser Ala Phe Cys Ile Phe Ser Val Leu Phe Thr Leu
327      420      425      430
329 Phe Cys Val Pro Glu Thr Lys Gly Lys Thr Leu Glu Gln Ile Thr Ala
330      435      440      445
332 His Phe Glu Gly Arg
333      450
336 <210> SEQ ID NO: 8
338 <211> LENGTH: 53
340 <212> TYPE: PRT
342 <213> ORGANISM: homo sapiens
344 <220> FEATURE:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/516,493

DATE: 07/18/2001

TIME: 10:37:01

Input Set : A:\96700-613.2.ST25.txt

Output Set: N:\CRF3\07182001\I516493.raw

```

346 <221> NAME/KEY: varsplic
348 <222> LOCATION: (1)..(53)
350 <223> OTHER INFORMATION: possible alternate carboxy terminus of predicted
351     amino acid sequence of human GLUTx protein
353 <400> SEQUENCE: 8
355 Gln Pro Leu Thr Arg Gly Trp Ser Lys Pro Val Thr Pro Ser Trp Ala
356 1             5             10             15
358 Gln Ala Gln Ser Pro Cys Leu Pro Gln Gly Ser Gln Asn Pro Ala Pro
359             20             25             30
361 Trp Ser Leu Gly Leu Gln Gly Pro Ser Phe Leu Ser Cys Ser Leu Gln
362             35             40             45
364 Pro Met Thr Arg Gly
365             50
368 <210> SEQ ID NO: 9
370 <211> LENGTH: 1037
372 <212> TYPE: DNA
374 <213> ORGANISM: rat
376 <220> FEATURE:
378 <221> NAME/KEY: unsure
380 <222> LOCATION: (606)..(609)
382 <223> OTHER INFORMATION: 'n' may be any one of a, t, c, or g
384 <400> SEQUENCE: 9
386 tggcggccgc tctagaacta gtggatcccc cgggctgcag gaattcggca cgagctggtg      60
388 cccatctccg cagagcctgc tgatgttcac ctggggctgg cctggctggc tgtaggcagc      120
390 atgtgcctct tcacgcctgg ttttgcagta ggctggggac ccatcccctg gctcctcatg      180
392 tcagagatct tccctctgca catcaagggt gtggctaccg gcgtctgtgt cctcaccaac      240
394 tggttcatgg cctttctggt gaccaaagag tttaacagca tcatggagat cctcagaccc      300
396 tacggcgccct tctggctcac cgctgccttc tgtatcctca gcgtcctttt cacgctcacc      360
398 tttgtccctg agactaaagg caggactctg gaacaaatca cagcccatth gagggacggg      420
400 gacggacccc tttctgtgac tggcagccct gagctgagct ggcttcgggt ttcaaaagga      480
402 gtggagtggc ctcaagtacc acagtttgag cccagggggc ccctgactcc tcagatttcc      540
404 gggccagctt tgtccagatc tcaaccaga ttccacacca tgagcttcac cagattctga      600
406 ggctctgna gcctgctgca cacacagcac atttgcgggc tcctggctct agtgtctctg 660
408 ctgggcatct ttgggggtgct tggtcctaag caactgccca tacctcactt gactggggga      720
410 tgagaaaagg acttagccac ataagatttg ggctcagaaa caaggtcagg tgagtccagg      780
412 aagaaaagag aatggttctt gtcttgtcaa ccaagtcctt ctcaagtgc caaagacctc      840
414 cggattcacc ttgggggttag ccagcttacc catcacttac aggttctctc caactctcag      900
416 ctggtctcag tgtcctggat cattagtcac caggtcttgt tgagtttcag aaaaataaaa      960
418 ggcctctttc cgttcaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa     1020
420 aaactcgagg gggggcc
423 <210> SEQ ID NO: 10
425 <211> LENGTH: 165
427 <212> TYPE: PRT
429 <213> ORGANISM: rat
431 <400> SEQUENCE: 10
433 Trp Arg Pro Leu Glx Asn Glx Trp Ile Pro Arg Ala Ala Gly Ile Arg
434 1             5             10             15
436 His Glu Leu Val Pro Ile Ser Ala Glu Pro Ala Asp Val His Leu Gly
437             20             25             30

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/516,493

DATE: 07/18/2001

TIME: 10:37:02

Input Set : A:\96700-613.2.ST25.txt

Output Set: N:\CRF3\07182001\I516493.raw

L:230 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6

L:406 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9